

Dealing with Slow File Retrieval

Category: Storage Best Practices

There are sometimes problems with commands on Lou, that should finish quickly, but end up taking a long time.

When you do an `ls` on Lou, you see all the files on disk that you've put there. However, most of the files are actually written to tape using SGI's Data Migration Facility (DMF).

One problem with DMF is that it does not deal well with retrieving one file at a time from a long list of files. If you do an `scp` with a list of files, UNIX feeds those files to DMF one at a time. This means that the tape(s) containing the files is getting constantly loaded and unloaded which is bad for the tape and tape drive, and also very slow. As the list of files gets longer (by use of "*" or moving a "tree" of files) the problem grows to where it can take hours to transfer a set of files that would only take a few minutes if they were on disk. When several people do file transfers at once that retrieve files one at a time, it can tie the system in knots.

Optimizing File Retrieval

DMF let you fetch files to disk as a group with the `dmget` command. The tape is read once and gets all the requested files in a single pass. Essentially, give `dmget` the same list of files you are about to transfer, and when the `dmget` completes, then `scp/ftp/cp` the files as you had originally intended. Or you can put the `dmget` in the background and run your transfer while `dmget` is working. If any files are already on disk, `dmget` sees this and doesn't try to get them from tape.

There is also a `dmfind` command that let you walk a file tree to find offline files to give to `dmget`. Make very sure you are in the correct directory before running `dmfind`. Use the `pwd` command to determine your current directory.

Please check to make sure too much data isn't brought back online at once by using `du` with the `--apparent-size` option or by using `/usr/local/bin/dmfdu`.

```
lou% /usr/local/bin/dmfdu Foo
Foo
      13 MB regular          340 files
    1114 MB dual-state      1920 files
   74633 MB offline        2833 files
      13 MB small          340 files
   75761 MB total          5093 files
```

When transferring data between Lou and Columbia nodes use the /nobackup filesystems, instead of the Columbia NFS (slow) home directories.

File transfer rates vary depending on the load on the system and how many users are transferring files at the same time. Transferring files using **scp** between Lou and Pleiades on the /nobackup file system for files larger than 100 MB is typically between 30-120 MB/s using the 10-gigabit network interface.

Example 1:

```
lou% dmget *.data &  
lou% scp -qp *.data myhost.jpl.nasa.gov:/home/user/wherever
```

Example 2:

```
lou% dmfind /u/username/FY2000 -state OFL -print | dmget &  
lou% scp -rqp /u/username/FY2000 some_host:/nobackup/user1/wherever
```

You can see the state of a file by doing **dm~~l~~s -l** instead of **ls -l**. For more information on using DMF, please look at: [Data Migration Facility DMF Commands](#).

Maximum Amount of Data to Retrieve Online

The online disk space for Lou1 and Lou2 is much, much less than its tape storage capacity, and it is impossible to retrieve all files to online storage at the same time. If Shift is used for file transfers, it will automatically ensure that files on Lou are retrieved in batches and released afterwards so there is no need to manually split up the transfer. If Shift is not used, however, then before retrieving a large amount of data, you should check that there is enough online space for it. The **df** command shows the amount of free space in a filesystem. The Lou script **dmfdu** reports how much total (online and offline) data is in a directory. To use this script, simply **cd** into the directory you want to know total amount of data for all the files in the current directory and execute the script.

If you would like to know the total amount of data under your home directory on Lou, you need to first find out if your account is under s1h or s2a-s2e. Assuming you are under s1h, you can then use **dmfdu /s1h/your_userid** to find the total amount. Another alternative is to simply **cd** to your home directory and use **dmfdu ***, which will show use for each file or directory.

Lou1 and Lou2's archive filesystems are between 60 TB and 450 TB in size, but the available space typically floats between 10% to 30%. In Example 3, 29% of space is unused. It is best to retrieve at most 10 TB at a time. Do what you need to with those files (**scp**, **edit**, **compile**, etc), then release (**dmpout -r**) the space, and then retrieve the next group of files, use them, then release the space, etc. For Example 3, retrieve one

directory's data from tape, copy the data to remote host then release the data blocks, before retrieving more data from tape.

Example 3:

```
lou% df -lh .
Filesystem              Size  Used Avail Use% Mounted on
/dev/cxvm/sfa2-s2l      228T  196T   32T   86% /lou/s2l

lou% dmfd du project1 project2
project1
      2 MB regular                214 files
     13 MB dual-state              1 files
2229603 MB offline               101 files
      2 MB small                  214 files
2229606 MB total                  315 files

project2
      7 MB regular                245 files
    4661 MB dual-state             32 files
2218999 MB offline                 59 files
      7 MB small                  245 files
2223668 MB total                  336 files

lou% cd project1

lou% dmfind . -state OFL -print | dmget &

lou% scp -rp /u/username/project1 remote_host:/nobackup/username

(Verify that the data has successfully transferred)

lou% dmfind . -state DUL -print | dmput -rw

lou% df -lh .

lou% cd ../project2

lou% dmfind . -state OFL -print | dmget &

lou% scp -rpq /u/username/project2 remote_host:/nobackp/username

lou% dmfind . -state DUL -print | dmput -rw
```

Article ID: 150

Last updated: 05 Feb, 2013

Data Storage & Transfer -> Storage Best Practices -> Dealing with Slow File Retrieval

<http://www.nas.nasa.gov/hecc/support/kb/entry/150/?ajax=1>